COMP 190-088: Systems Performance Analysis

System Modeling - II

Jasleen Kaur

Department of Computer Science
The University of North Carolina at Chapel Hill

Spring 2005

Kendall Notation

- A/S/m/B/K/SD
  - A: inter-arrival time distribution
  - S: service time distribution
  - m: number of servers
  - B: number of buffers
  - K: population size
  - SD: service discipline
- A and S generally denoted by one-letter symbols:
  - M: Exponential
  - G: General
  - D: Deterministic (constant service time)
- Bulk arrivals and bulk service:
  - Each arrival/service consists of a group of jobs
  - Bulk Poisson arrival denoted by $M^{[x]}$, where X is the group size

Notation Examples

- $M^{[x]} / G^{[x]}$ /3/20/1500/FCFS
  - Bulk Poisson arrivals
  - Bulk service with general service time distribution
  - 3 servers with 20 buffer slots
  - Population of 1500
  - FCFS service order
- Unless specified, assume infinite buffer, infinite population and FCFS
  - eg, G/G/1 implies $G/G/1/\infty/\infty$/FCFS

Quantities of Interest

- λ: arrival rate
- μ: service rate
- N(t): number of requests in the system at time t
- Nq(t): number of requests waiting in the queue at time t
- Ns(t): number of requests in service at time t
- R: Response time for a request
- W: Waiting time for a request
- S: service time for a request
- U: system utilization

Rules for All Queues (G/G/m)

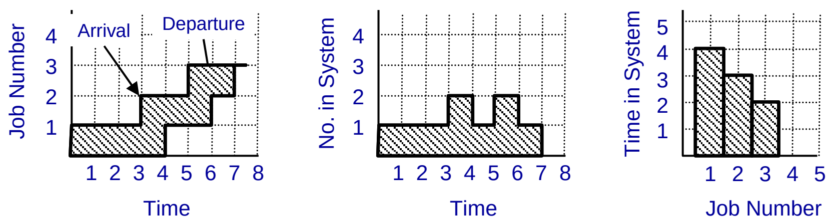
- Stability Condition:
  - System is unstable if number of jobs in system grows continuously and becomes infinite
  - For stability: $\lambda < m*\mu$
  - Condition not necessary in finite population and finite buffer systems
- Number in system (N), number in queue (Nq), number in service (Ns)
  - N = Nq + Ns
  - E[N] = E[Nq] + E[Ns]
  - If service rate is independent of queue size,
    - Cov(Nq,Ns) = 0
    - Var[n] = Var[Nq] + Var[Ns]
- Time in system (R), time in queue (W), service time (S)
  - R = W + S

Little's Law

- If number of jobs entering system is equal to the number completing service, then: $E[N] = \lambda * E[R]$
  - Mean number of jobs in system = arrival rate * mean response time
  - If jobs are not lost due to buffer overflow
  - If no jobs are created or lost inside the system
- Applies to all sub-systems in which number of jobs entering and leaving the system is the same
  - Mean number of jobs in queue = arrival rate * mean waiting time
  - Mean number of jobs in service = arrival rate * mean service time
- In finite-buffer systems
  - Applies to system if arrival rate adjusted to exclude lost jobs

Proof For Little's Law

- Suppose all arrivals and departures in time T are logged
  - If T is large, # of arrivals = # of departures = N
  - Arrival rate = N/T
- 3 ways to plot the logged data:



- All three areas are equal = J
- Mean time spent in system = J/N
- Mean number in system = J/T = N/T * J/N = arrival rate * mean time in system

Types of Stochastic Processes

- Discrete-state vs. continuous-state
- Markov processes
- Birth-death processes
- Poisson processes

Discrete-state Vs. Continuous-state

- Discrete-state:
  - If number of values the process state can take is finite or countable
  - eg, number of jobs in system
- Continuous-state:
  - State can take uncountably infinite values
  - eg, waiting time in queue